Kapitel 1

Kombinatorische Optimierung

Vielen Problemen der Optimierung und der Informatik liegen diskrete Strukturen zugrunde, über die für lineare oder nichtlineare Funktionen ein Optimum bestimmt werden soll. In Projekten der kombinatorischen Optimierung werden diese Probleme zunächst hinsichtlich ihrer Komplexität charakterisiert. Danach wird versucht, effiziente Algorithmen zu entwickeln, wobei – abhängig von der Komplexität des Problems – exakte, approximative oder randomisierte Algorithmen relevant werden.

Ihre herausragende Bedeutung in den letzten Jahren verdankt die kombinatorische Optimierung vor allem der Verschiedenartig- und Wichtigkeit vieler ihrer Anwendungsgebiete. Viele Entscheidungssituationen, wie sie etwa in Industrie und Wirtschaft auftreten, lassen sich auf natürliche Weise als ganzzahlig-lineare Optimierungsprobleme über endlichen, allerdings meist astronomisch großen Mengen formulieren. Tiefergehendes theoretisches Verständnis der zugrundeliegenden mathematischen Struktur des Lösungsraumes ermöglicht vielfach die Entwicklung beweisbar „guter“ und in der Praxis erfolgreicher Algorithmen.

Dieses Kapitel veranschaulicht sowohl beispielhaft anwendungsbezogene Aspekte kombinatorischer Techniken als auch theoretische Forschungsergebnisse unserer Arbeitsgruppe.

1.1 Optimierungsprobleme bei Fluggesellschaften

Durch die Globalisierung der Märkte und Internationalisierung von Unternehmen wird der effiziente Einsatz personeller und materieller Ressourcen zu einem wesentlichen Faktor für die Konkurrenzfähigkeit von Transport- und Logistikunternehmen. Bei Fluggesellschaften spielt hierbei die Flottenzuweisung eine wichtige Rolle. Sie bildet die Basis für viele weitere Aktivitäten wie Marketing und Preispolitik. Aber auch die Personaleinsatzplanung bietet aufgrund der hohen Aufwendungen für das Flugpersonal ein hohes Optimierungspotential.

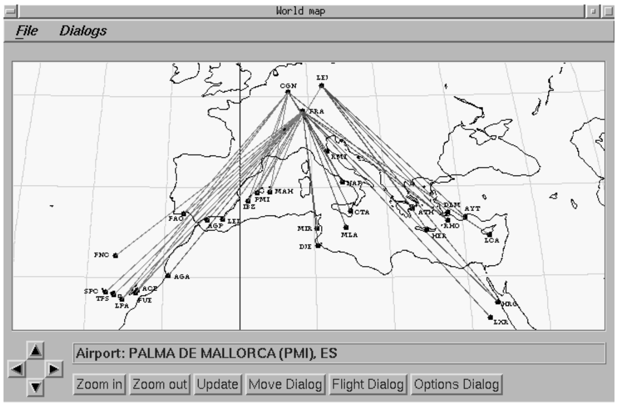

Abbildung 1.1: Ausschnitt aus einem Flugplan einer Charterfluggesellschaft

Die ökonomische Wertigkeit der zu treffenden Entscheidungen ist häufig für den kommerziellen Erfolg ausschlaggebend, insbesondere angesichts der hohen Produktionskosten am „Standort Deutschland“. Bei großen Fluggesellschaften wie der Lufthansa reicht die Ertragswertigkeit der in der Flugplanstruktur und im Ertragsmanagement festzulegenden Erfolgsfaktoren an etwa eine Milliarde DM pro Jahr heran. Im Bereich des Fleet-Assignments oder des Crew-Schedulings geht es um mehrere 100 Millionen DM jährlich sowie um die immateriellen Vorzüge einer verbesserten Personaleinsatzplanung, die schließlich die Qualität des Arbeitslebens des fliegenden Personals wesentlich prägt. Dies bedeutet, daß schon Verbesserungen von mehreren Promille oder gar Prozent wesentlich zur Wettbewerbsfähigkeit und Servicequalität beitragen.

Der Planungsablauf bei Fluggesellschaften

Bei Fluggesellschaften kann der Planungsablauf für den Passage-Bereich in die folgenden Schritte unterteilt werden. Während der *Produktplanungsphase* wird die Nachfrage geschätzt und aufbauend auf diesen Informationen werden die angebotenen Flugstrecken festgelegt. In der *Produktionsplanung*, die immer noch ca. ein halbes Jahr vor dem eigentlichen Ausführungstermin stattfindet, werden die genauen Abflugszeitpunkte und die zu benutzenden Flugzeugtypen (*Fleet-Assignment*) bestimmt. Erst wenige Wochen vor dem geplanten Zeitraum werden während der *Ressourcenallokation* den einzelnen Flugzeugen die Umläufe (d. h. die zu bedienenden Flugstrecken) zugewiesen (*Tail-Assignment*) und jeder einzelne Flug mit Flugpersonal (*Crew-Assignment* und *-Rostering*) versehen. Häufig muß die Fluggesellschaft während der Ausführung reagieren. In der *Operation-Time-Phase* werden z. B. aufgrund von witterungs- oder materialbedingten Störungen Flugzeugumläufe angepaßt. Diese Schritte werden nicht bei jeder Planung vollständig neu durchgeführt, sondern Änderungen werden schrittweise eingebracht.

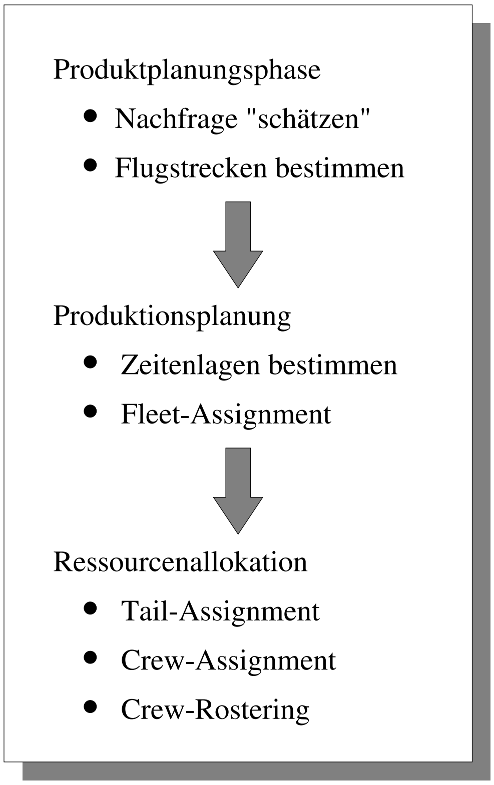

Abbildung 1.2: Ablauf des Planungsprozesses bei Fluggesellschaften

Am ZAIK/ZPR wird intensiv an Problemen aus dem Bereich der Flugplanoptimierung gearbeitet. Dabei bilden die Probleme Fleet-Assignment und Crew-Assignment die Forschungsschwerpunkte.

1.1.1 Fleet-Assignment

Eingabe für dieses Teilproblem ist der Flugplan, der in den vorhergehenden Planungsschritten bestimmt wurde. Dieser enthält genaue Informationen über die anzubietenden Flugstrecken und geplanten Abflugszeiten. Zu den weiteren Eingaben zählen Flugzeugeinsatzkosten (wie Treibstoffkosten) und detaillierte Marktinformationen (wie Passagiernachfragen und -erlöse auf Start-/Ziel-Beziehungen).

Ziel dieser Optimierung ist die Zuweisung der zur Verfügung stehenden Flugzeugflotten zu den Flügen, so daß der Gesamtgewinn maximiert wird. Dabei sind vielfältige operationelle Nebenbedingungen zu beachten. Neben den offensichtlichen Vorgaben wie Flugzeugreichweiten und Start-/Landeerlaubnis müssen z. B. auch die Flottengrößen und die Periodizität der Flottenzuweisung eingehalten und Wartungszeiten berücksichtigt werden.

Zur Lösung dieser komplexen Optimierungsprobleme verwenden wir von uns entwickelte lokale Suchverfahren (wie Simulated Annealing), duale Verfahren (wie Lagrange-Relaxierung) sowie Methoden, die auf Linearer Programmierung und Netzwerkalgorithmen basieren.

Ergebnisse

Die entwickelten und implementierten Algorithmen wurden in ein Programmpaket integriert, mit dessen Hilfe Flugplaner sowohl mittelfristige Flugplanänderungen durchführen als auch langfristig hypothetische Szenarien auf ihre Wirtschaftlichkeit prüfen können. Durch eine Workbench mit grafischer Benutzeroberfläche ist es möglich, die Algorithmen zu starten und zu steuern, Änderungen an den Plänen per Hand vorzunehmen und die Ergebnisse auf mehrere verschiedene Arten anzuzeigen (geographische Sicht, als Gantt-Diagramm, Listen, etc.).

Zur Entwicklung und zum Vergleich der Algorithmen liegen uns mehrere „Real-World“-Datensätze vor, die uns von der Lufthansa zur Verfügung gestellt wurden: kleine „interkontinentale“ Datensätze mit 300 bis 3.500 Flugstrecken, bis zu fünf verschiedenen Flugzeugtypen und bis zu 50.000

Reiserouten, aber auch sehr große „europäische Wochenflugpläne“ mit 7.500 Flugstrecken, 35 verschiedenen Flugzeugtypen und 500.000 Reiserouten.

Unter Verwendung der oben angesprochenen Algorithmen ist es uns möglich, sämtliche Datensätze zu lösen und die (annähernde) Optimalität nachzuweisen. So ermöglicht z. B. unsere Lösung des größten Datensatzes eine Einsparung von ungefähr 4% gegenüber den verwendeten Flotteneinsatzplänen. Laufzeiten von weniger als 20 Stunden auf Computern der Workstationklasse liegen dabei in einem für die Langfristplanung gut vertretbaren Rahmen.

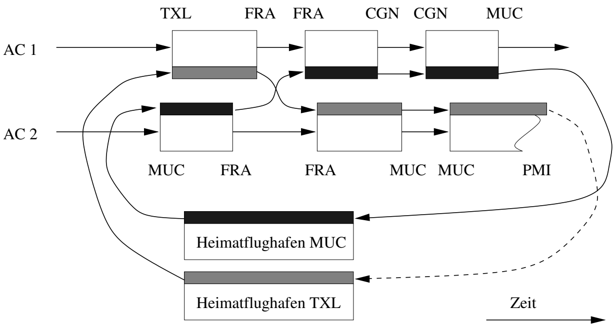

Abbildung 1.3: Ein einfaches Crew-Scheduling Beispiel mit zwei Flugzeugen

1.1.2 Crew-Assignment

Ziel dieses Planungsschrittes ist es, durch geschickte Zuweisung der Crews zu den angebotenen Flügen entstehende Kosten (das sind im wesentlichen Hotelkosten, Kosten durch Transport der Crews von und zu ihrem Einsatzort und unnötige Wartezeiten) zu minimieren und so eine Gewinnsteigerung des Unternehmens zu ermöglichen.

Problemstellung

Gegeben sind in diesem Problem die Flugzeugumläufe, d. h. die Folge von Flügen, die von jedem einzelnen Flugzeug geflogen werden; diese werden in den Optimierungsschritten Fleet-Assignment und Tail-Assignment bestimmt.

Abbildung 1.3 zeigt ein Beispiel für ein einfaches Crew-Scheduling-Problem. In dem dargestellten Ausschnitt bedient das Flugzeug AC1 die Flüge Berlin (TXL) → Frankfurt (FRA), Frankfurt → Köln (CGN) und Köln → München (MUC) und das Flugzeug AC2 die Flüge München → Frankfurt, Frankfurt → München und München → Palma (PMI). Eine Crew mit Heimatflughafen TXL begleitet den Flug TXL → FRA des Flugzeugs AC1, wechselt dort auf das Flugzeug AC2 und begleitet dieses auf den Flügen FRA → MUC, MUC → PMI und kommt am Ende des Dienstes wieder am Heimatflughafen an. Entsprechendes gilt für eine Crew mit Heimatflughafen MUC, die die restlichen Flüge begleitet.

Bei der Zuweisung der Crews zu den angebotenen Flügen müssen neben den offensichtlichen Nebenbedingungen (wie z. B. die Verfügbarkeit der Crews an den einzelnen Heimatflughäfen und Zulassungen für Flüge und Flugzeugtypen) auch sehr viel subtilere beachtet werden. Dazu zählen tarifrechtliche und sicherheitsrelevante Bestimmungen, die Beachtung von Reserven und die Berücksichtigung von Mitarbeiterwünschen. Insbesondere gilt es, die Beschränkung hinsichtlich der Flug-, Arbeits- und Ruhezeiten einzuhalten, die auf europäischer Ebene in den JAR-OPS[1] festgehalten sind. Darin werden die Dienstschichten in Folgen von Arbeitszeiten, die durch Pausezeiten verbunden werden, zerlegt. Zu den in JAR-OPS formulierten Regelungen zählen unter anderem

- die Beschränkung der Gesamtdauer einer Dienstschicht, der Gesamtdauer der Arbeitszeitblöcke und der darin enthaltenen Flugzeit,
- die Beschränkung jedes einzelnen Arbeitszeitblocks in Abhängigkeit von dem Zeitpunkt des ersten Starts und der Anzahl der enthaltenen Landungen, und
- Anforderungen an die Ruhezeiten: diese müssen jeweils eine gewisse Mindestlänge haben, deren Dauer von dem vorhergehenden Arbeitszeitblock abhängt.

Mathematische Modellierung und Lösungsansatz

Wir verwenden zur Lösung eine Set-Partitioning-Formulierung mit Nebenbedingungen, die als Lineares Programm modelliert wird: aus der Menge aller möglichen Dienstschichten muß eine Teilmenge ausgewählt werden, so daß, unter Beachtung von globalen Nebenbedingungen

[1] *Explanatory note covering the Joint Aviation Authorities flight and duty time limitations and rest requirement proposals*, Flight Time Limitations Study Group (FTLSG) of the Joint Aviation Authorities Operations Committee, 1995.

wie Kapazitäten an den Heimatflughäfen, jeder Dienst (=Flug) von genau einer Crew überdeckt wird.

Die Anzahl der möglichen Dienstschichten ist schon für kleine Probleme extrem groß. So enthält z. B. eine Probleminstanz mit nur 105 Flügen und sechs Heimatflughäfen über 1.350.000 verschiedene zulässige Dienstschichten, die während der Optimierung beachtet werden müssen.

Ein direkter Lösungsansatz ist daher bei praxisrelevanten Problemgrößen nicht möglich, so daß wir uns für den Einsatz der Spaltengenerierungstechnik entschieden haben. Beginnend mit einem vereinfachten Linearen Programm, das nur eine vielversprechende Auswahl aller Dienstschichten enthält, werden in einem iterativen Prozeß abwechselnd das Lineare Programm gelöst und unter Verwendung dieser Lösung dynamisch neue Dienstschichten generiert. Dieses Teilproblem, das sogenannte Subproblem, ist ein restringiertes Kürzeste-Wege-Problem. Mit Hilfe von „Ressourcen“ werden in einem Multi-Label-Algorithmus interessante neue Dienstschichten erzeugt, die alle tarifrechtlichen und sicherheitsrelevanten Nebenbedingungen einhalten.

Um ganzzahlige Lösungen zu erhalten, wird dieses Verfahren in einen Branch-&-Bound-Baum eingebunden. Als Branchingregel für den linken Teilbaum erzeugen wir „ $a \to b$ “: die Crews müssen immer beide Flüge a und b oder keinen dieser beiden Flüge bedienen. Entsprechend gilt „ $a \not\to b$ “ für den rechten Teilbaum, was besagt: die Crews müssen entweder a oder b, aber nicht beide oder keinen bedienen.

Ergebnisse

Unter Verwendung speziell für dieses Problem entwickelter Preprocessing-Techniken, Such- und Auswahlstrategien für den Branch-&-Bound-Prozeß und Heuristiken ist es mit der entstandenen Umgebung möglich, „Real-World“-Probleme zu lösen.

Von einer großen deutschen Fluggesellschaft wurden uns Datensätze mit bis zu 1.400 Flügen zur Verfügung gestellt. Bei Laufzeiten um ca. 20 min für den größten Datensatz können damit auf handelsüblichen PCs (Pentium-II 333 MHz) Dienstpläne erzeugt werden, die eine Einsparung von fast 3% gegenüber den verwendeten Dienstplänen ermöglichen. Die attraktive kurze Laufzeit erlaubt weiter den Einsatz in Decision-Support-Systemen.

1.1.3 Aktuelle Forschung: Generierung saisonaler Charterflugpläne

Aufbauend auf unseren Erfahrungen bei den zuvor genannten Optimierungsproblemen in der Flugplanung arbeiten wir zur Zeit an einem neuen Themenkomplex. Speziell Charterfluggesellschaften sehen sich dem Problem sehr saisonabhängiger Flugpläne gegenüber, so daß hier Informationen über den zeitlichen Verlauf des Passagieraufkommens stärkeren Einfluß auf die Flugplangestaltung haben muß als dies bei Linienfluggesellschaften der Fall ist.

Problemstellung

Eingabe für dieses Optimierungsproblem sind saison- und wochentagabhängige prognostizierte Nachfragen für alle angebotenen Flugstrecken. Diese bestehen aus den durch die enge Kopplung der Charterfluggesellschaft mit den Reiseveranstaltern garantierten Passagiernachfragen und den Daten der vergangenen Jahre. Abbildung 1.1 zeigt einen kleinen Ausschnitt aus dem Angebotsnetzwerk. Eingezeichnet sind hier drei deutsche Flughäfen (Köln (CGN), Frankfurt (FRA) und Leipzig (LEJ)) und zahlreiche Urlaubsziele wie Fuerteventura (FUE), Djerba (DJE) und Hurghada (HRG).

Das Problem besteht nun darin, die zur Verfügung stehenden Flugzeuge der Charterfluggesellschaft so einzuplanen, daß ein unter Profitabilitätsgesichtspunkten optimaler Flugplan entsteht. Der Gewinn eines Flugplanes errechnet sich dabei sowohl aus den Fix- und Betriebskosten der Flugzeuge (wobei nicht eingesetzte Flugzeuge an andere Fluggesellschaften vermietet oder Flugzeuge angemietet werden können) als auch aus den passagierabhängigen Kosten und Erträgen.

Weiter müssen operationelle Nebenbedingungen wie Flugzeugreichweiten, Flughafenkapazitäten und -öffnungszeiten sowie Mindestbodenzeiten (zum Auftanken und Catering) beachtet werden.

Mathematische Modellierung und Lösungsansatz

Entsprechend unseren guten Erfahrungen mit den auf Linearer Programmierung basierten Ansätzen für das Fleet- und Crew-Assignment-Problem verwenden wir auch hier zur Lösung eine Set-Partitioning Formulierung, wobei die

Berechnung der mitgenommenen Passagiere über ein integriertes Set-Packing Subproblem geschieht.

Aufgrund der geringen Festlegung durch globale Restriktionen ist die Anzahl der möglichen Flugzeugrouten extrem hoch, weil grundsätzlich, abgesehen von Reichweitebedingungen, fast jedes Flugzeug jede Strecke bedienen könnte. Als Konsequenz benutzen wir auch hier einen dynamischen Spaltengenerierungsansatz. Das Subproblem zur Erzeugung neuer Flugzeugumläufe läßt sich dabei wieder als restringiertes Kürzeste Wege-Problem formulieren und mit einem Multi-Label-Algorithmus lösen.

Im Gegensatz zu unseren Erfahrungen mit den Fleet Assignment und den Crew Scheduling Problemen stellt die LP-Relaxierung bei dieser neuen Problemklasse eine schlechte Approximation des Raumes aller möglichen ganzzahligen Lösungen dar. Zur Zeit liegen die nachweisbaren Gütegarantien um 10-15%. Um diese zu verbessern und auch hier Optimalität zu beweisen, arbeiten wir zur Zeit an einem kombinierten Branch-&-Cut-&-Price Ansatz, bei dem das „Gap“ durch zusätzliche Schnittebenen verkleinert wird.

Ergebnisse

Unter Verwendung des oben vorgestellten kombinierten Branch-&-Bound- und Spaltengenerierungs-Ansatzes ist es möglich, Problemgrößen, wie sie bei der „Condor“, einer der größten deutschen Charterfluggesellschaften, auftreten, zu bearbeiten – diese enthalten bis zu 20 deutsche und 60 Zielflughäfen bei bis zu 60 eigenen Flugzeugen. Wir erhalten dabei Laufzeiten, die im Bereich unter einer Stunde liegen, Flugpläne, die um 1-2% besser sind als die, die von den strategischen Planern bei Condor erstellt worden sind. Leider können wir jedoch zur Zeit keine verläßlichen Aussagen darüber machen, ob wir mit unseren Plänen schon das Optimierungspotential ausgereizt haben. Da, wie oben erwähnt, das „Gap“ sehr groß ist, wird dieses Problem durch intensive Forschung im Bereich Schnittebenenverfahren zur Zeit von uns bearbeitet.

1.1.4 Übertragung auf ähnliche Problemstellungen

Die entwickelten Algorithmen können leicht auf verwandte Probleme übertragen werden — z. B. die Planungsprobleme der Verkehrsbetriebe.

Seit am 1. Januar 1996, also zwei Jahre nach der Bahnreform, eine EU-Richtlinie von Bund und Ländern in deutsches Recht umgesetzt wurde, müssen die öffentlichen Hände die Nahverkehrsleistungen bei den Verkehrsunternehmen einkaufen. Dabei stellt, bei fortschreitender Privatisierung und zunehmendem Wettbewerb auf dem ÖPNV-Markt, der effiziente Ressourceneinsatz einen immer wichtigeren Faktor dar.

Zur Planung der Omnibusse und der Fahrer könnte eine optimierende Planungsumgebung den entscheidenden Wettbewerbsvorteil gegenüber anderen Anbietern schaffen.

1.2 Ford – Wir tun auch was

„Kundenzufriedenheit durch kürzere Lieferzeiten“, mit diesem Ziel startete die Firma Ford das Projekt „Order to Delivery“ zur Neustrukturierung der gesamten EDV-technischen Auftragsbearbeitung. Angefangen von der Prognose zukünftiger Auftragsbestände über die Erfassung bis hin zur Abwicklung der Aufträge wurden alle Schritte neu konzipiert. Für die *Produktionssegmentierung*, bei der jeder Bestellung ein optimales Produktionsdatum zugeordnet wird, wurde das später umgesetzte Konzept von unserer Arbeitsgruppe entwickelt. Bereits in der Vergangenheit gab es bei Ford ein System, das Fahrzeuge so in die Werke einsteuerte, daß verschiedene Verkaufsdistrikte und Händler regelmäßig mit Neuwagen beliefert werden konnten. Nicht berücksichtigt wurden dabei jedoch viele der Restriktionen, die sich aus der Produktion in den Werken und insbesondere bei den Zulieferern ergaben. Hierin lag noch ein großes Optimierungspotential, das durch den Einsatz von linearer Programmierung erfolgreich erschlossen werden konnte.

Mehrere Monate vor dem eigentlichen Produktionstermin werden von den Landesorganisationen Prognosen für den zukünftigen Auftragsbestand erstellt. Auf Basis dieser Daten wird dann eine Ressourcenzuordnung zu den einzelnen Märkten vorgenommen und es werden Bestellungen bei den Zulieferern abgegeben. Ressourcen sind zum Beispiel einzelne Bauteile, wie Fahrgestelle, Motoren oder Sonnendächer. Die realen Auftragsdaten gehen erst wenige Wochen vor dem Produktionstermin ein. Sie bestehen aus Kombinationen der einzelnen Bauteile, das heißt, es wird ein Fiesta mit Sonnendach und anderen Ausstattungsmerkmalen bestellt. Ziel der Optimierung ist nun neben der Auslastung der Produktionskapazitäten eine möglichst hohe Übereinstimmung der Prognosedaten mit den Ergebnissen der realen Produktionsplanung. Dabei muß garantiert sein, daß bei keinem Produktionsschritt die vorhandenen Kapazitäten überlastet werden.

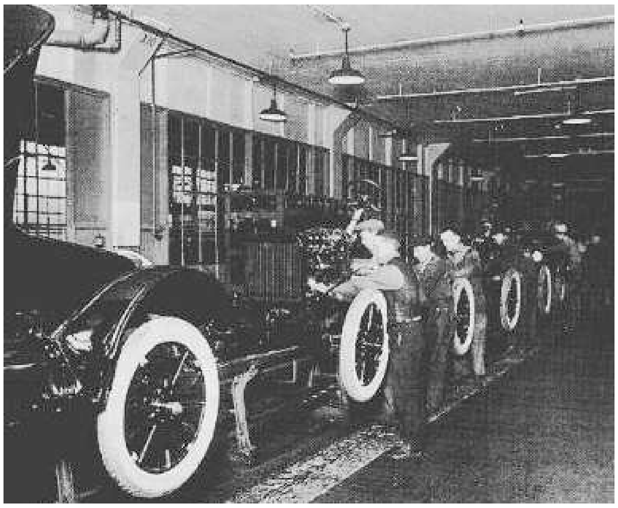

Abbildung 1.4: Model T-Produktion bei Ford zu Beginn des Jahrhunderts.

Für jeden Produktionstag muß also eine Menge von Aufträgen gefunden werden, so daß ein geeigneter Bauteilmix entsteht. Die produktionsbedingten Restriktionen lassen sich in der Regel durch obere Schranken für einzelne Komponenten sinnvoll abbilden. Daraus wurde ein lineares Modell entwickelt, dessen Zielfunktion die vorgegebenen Optimierungskriterien widerspiegelt. Mit diesem Ansatz kann für jeweils einen Monat eine gute Segmentierung errechnet werden. Nachträglich eingehende Bestellungen aber auch die unvermeidlichen Störungen des Produktionsablaufes machen es allerdings notwendig, in regelmäßigen Abständen neu zu planen. So kann es z. B. immer zu Lieferengpässen bei Zulieferern oder auch zu Streiks unter den LKW-Fahrern kommen. In diesen Situationen soll der neu berechnete Plan sich möglichst wenig von dem ursprünglichen Plan unterscheiden. Die Händler sind so in der Lage ihren Kunden frühzeitig den voraussichtlichen Auslieferungstermin für das von ihnen bestellte Fahrzeug mitzuteilen. Diese Forderung nach Planungsstabilität macht es notwendig, den Ansatz zu erweitern, es ist aber möglich, auch dies als lineares Programm zu modellieren.

Das Konzept wurde zunächst an unserem Institut testweise implementiert und anhand von Beispieldaten für die europäische Fiestaproduktion auf seine Tauglichkeit überprüft. Es zeigte sich, daß mit dem neuen Modell die Planungssicherheit deutlich erhöht werden konnte, was sich in erheblichen Kostenersparnissen niederschlägt. Daneben bietet die Verwendung linearer Programme viele praktische Vorteile, da zu ihrer Lösung etablierte Standardsoftware zur Verfügung steht. Es konnte insbesondere durch den Einsatz von Mixed-Integer-Solvern erreicht werden, daß in fast allen Fällen optimale ganzzahlige Produktionsvorgaben errechnet werden können. Inzwischen wird das auf dem neuen Ansatz basierende Planungstool für nahezu alle europäischen Produktionsstätten mit Erfolg eingesetzt.

In der praktischen Arbeit mit dem Segmentierungstool zeigte sich, daß die für den Datentransfer und die Datenkonvertierung notwendige CPU-Zeit merklich größer ist, als die für die eigentliche Planung benötigte Laufzeit. Auch hier gab es die Möglichkeit, durch den Einsatz mathematischer Methoden deutliche Zeitersparnisse zu erreichen. Die einzelnen Ausstattungsmerkmale der Bestellungen sind durch logische Ausdrücke codiert, die effizient ausgewertet werden müssen. Hierzu haben wir ein Programm entwickelt, das im kommenden Jahr in den Prozess integriert werden soll.

1.3 Tourenplanung

Qualitativ hochwertige Logistikdienstleistungen können nur durch ständige Innovationsbereitschaft erbracht werden. Deshalb gewinnt computergestützte Tourenplanung bei Transportunternehmen immer weiter an Akzeptanz. Die Komplexität und Verschiedenartigkeit der dabei an die Optimierungsalgorithmen gestellten praktischen Aufgaben verlangen oft völlig neue Herangehensweisen. Im Rahmen einer Promotion wurde in unserer Gruppe die Arbeit an einem Verfahren fortgeführt, das auf dem *Cluster First – Route Second* Ansatz beruht. Dieser Algorithmus entwickelte sich aus speziellen Anforderungen des Anwenders an seine Tourenpläne.

Die Vernetzung verschiedener Verkehrsträger und damit auch das Verladen von Gütern während des Transports gewinnt immer größere Bedeutung. Eine optimale Nutzung solcher Logistikketten kann nur erreicht werden, wenn Planungswerkzeuge zur Verfügung stehen, die solche neuartigen Möglichkeiten integrieren. Mit seinen Partnern treibt das ZAIK/ZPR die Entwicklung innovativer Algorithmen für die intermodale Tourenplanung voran.

Weiterhin wurde im Berichtszeitraum von unserer Gruppe eine Fahrzeugeinsatzplanung in das Speditionsprogramm „ProfiTour" des Kooperationspartners Profi.S integriert. Damit sind die Algorithmen in der Lage, u. a. den Wiedereinsatz von LKW mit zu berücksichtigen. Dies ist ein Beispiel für Anforderungen an Distributionstouren, die beim praktischen Einsatz von Optimierungswerkzeugen unerläßlich sind.

In jüngster Zeit wurde unsere Forschung im Bereich Tourenplanung auch auf das Branch-&-Cut Verfahren ausgedehnt, mit dessen Hilfe sich Instanzen mit bis zu 70 Kunden exakt optimal lösen lassen.

1.3.1 Multimodaler Verkehr

Die weiter wachsenden Mobilitätsbedürfnisse unserer Gesellschaft manifestieren sich in einer immer höheren Auslastung der Infrastruktur für den Individualverkehr. So nahmen die Transportaktivitäten in der EU in den Jahren 1970 - 1990 um mehr als 85 % zu, der Güterverkehr sogar um mehr als 100 %. Aus diesen Gründen wird dringend nach neuen Strategien zur besseren Ausnutzung der vorhandenen Verkehrsinfrastruktur gesucht. Besonderes Augenmerk erhalten derzeit Ansätze, die einen effizienteren Einsatz der vorhandenen Ressourcen durch Vernetzung der verschiedenen Verkehrsträger erbringen sollen. So erhalten Transportunternehmen durch den Bau von intermodalen Frachtterminals die Möglichkeit, Waren während des Transports umzuladen. Die sich hieraus ergebenden gewaltigen Einsparungspotentiale, können allerdings nur durch einen erhöhten Planungsaufwand genutzt werden, da bei der Routenwahl nun eine größere Zahl an Möglichkeiten zur Verfügung steht. Für solche Tourenplanungsprobleme stehen zur Zeit keine marktreifen Planungswerkzeuge zur Verfügung. Das ZAIK/ZPR beschäftigt sich intensiv mit der Tourenplanung für den intermodalen Gütertransport. Es wird untersucht, wie sich eine Effizienzsteigerung durch den Einsatz intelligenter Umladestrategien sowie geschickter Dislozierung von Umladestationen erreichen läßt. Dies umfaßt so verschiedenartige Fragestellungen wie die Untersuchung intermodaler Ansätze zur Güterverkehrsplanung, die Verlagerung und Vernetzung des Verkehrs in den Ballungsräumen (Citylogistik) oder die Verwendung von Sammeldepots durch Branchen wie Paketdienste oder Speditionen. Es wurden in diesem Bereich Verbesserungalgorithmen entwickelt und diese werden zur Zeit an Beispielen aus der Realität getestet. Diese Forschung wird durch die Afried Krupp von Bohlen und Halbach-Stiftung gefördert. Eine Diplomarbeit zu diesem Thema wurde auch durch die Firma GEDAS vergeben und von unserer Arbeitsgruppe betreut.

1.3.2 Vehicle Routing Probleme

Das ZAIK/ZPR kann bereits auf eine langjährige Erfahrung im Bereich der Distributionstourenplanung zurückblicken. Häufig bereiten Nebenbedingungen, die nicht genau quantifiziert werden können, große Schwierigkeiten. Ein Unternehmen hatte den Wunsch, daß die Kunden einer Tour räumlich nahe beieinander liegen sollten, um die Ortskenntnis der Fahrer möglichst gut zu nutzen. Wir wählten daher die zur Lösung von Vehicle Routing Problemen (VRP) gängige Vorgehensweise des „Cluster first - Route second". Hierbei werden die zu beliefernden Kunden zunächst zu Clustern zusammengefaßt, um darauf aufbauend die Reihenfolge der Kunden eines Clusters, also die eigentlichen Touren bestimmen zu können.

Das uns damit gestellte Problem der Bestimmung von regional begrenzten Kundenclustern formulierten wir mathematisch als Baumzerlegungsproblem unter Nebenbedingungen. Ein Baum $T = (V, E)$ mit Knotengewichten $w : V \Rightarrow \mathbb{N}^q$ wird durch Entfernen einer Kantenmenge $S \subseteq E$ in Teilbäume zerlegt. Ziel ist es, eine minimale Menge S zu finden, so daß für die Summen der Knotengewichte in den einzelnen Teilbäumen T_i gilt: $\sum_{v \in V(T_i)} w(v) \leq R$ bei vorgegebener oberer Schranke $R \in \mathbb{N}^q$.

Komplexitätstheoretische Resultate

Erste Resultate waren ein allgemeiner NP-Vollständigkeitsbeweis und ein heuristischer Greedyansatz, der als Verallgemeinerung des von Kundu und Misra vorgeschlagenen Verfahrens zur Zerlegung von Bäumen mit einer Gewichtsfunktion aufgefaßt werden kann. Detailliertere Auseinandersetzungen mit dem Problem, den Praxisdaten und den NP-Vollständigkeitsbeweisen ergaben eine Unterscheidung der Baumzerlegungsprobleme sowohl nach dem Knotengrad der untersuchten Bäume als auch nach der Anzahl der Gewichtsfunktionen. Unter diesen Gesichtspunkten konnten wir eine vollständige Klassifizierung des

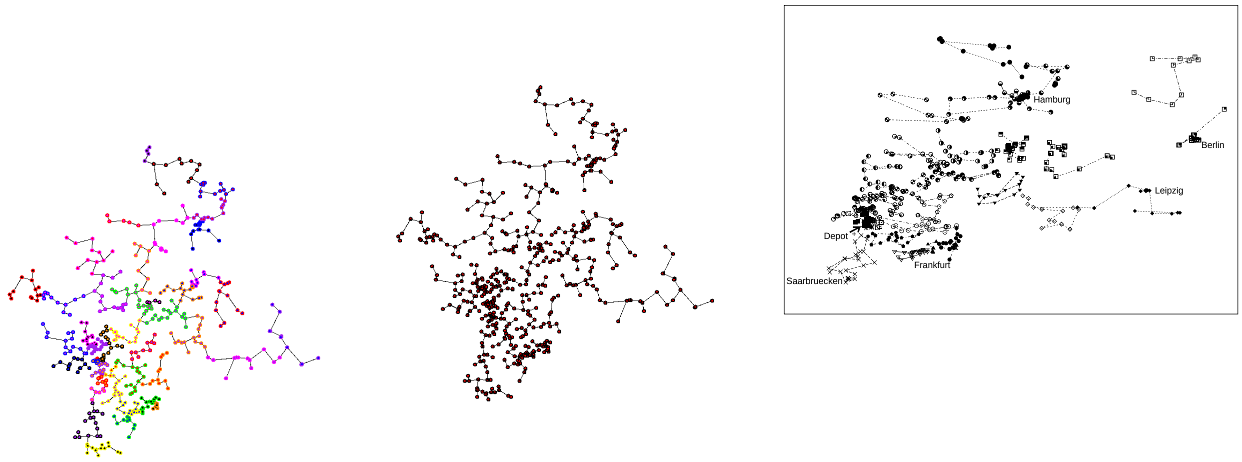

Abbildung 1.5: Cluster–First Route–Second

Baumzerlegungsproblems unter Nebenbedingungen angeben:

NP-vollständig im strengen Sinne ist das Baumzerlegungsproblem für Instanzen, bei denen die Anzahl der Gewichtsfunktionen mit der Problemgröße wächst. Bei Bäumen mit mindestens zwei Gewichtsfunktionen und unbeschränktem Knotengrad wird die Fragestellung zu einem Number-Problem. Es gelang uns hier, das Knapsack-Problem auf diese Teilklasse zu reduzieren und durch Angabe eines dynamischen Programmes die Pseudopolynomialität nachzuweisen. Bei Probleminstanzen mit einer Gewichtsfunktion ist der von Kundu und Misra beschriebene lineare Greedy-Algorithmus optimal. Für den bei der Clusterung auftretenden praxisrelevanten Fall der Zerlegung knotengradbeschränkter Bäume mit einer beschränkten Anzahl von Gewichtsfunktionen konnten wir ein exaktes polynomielles Verfahren entwickeln, das auf den Ideen des Greedy-Ansatzes aufbaut.

Clustern durch Baumzerlegungen

Eine regionale Zusammenstellung von Kunden zu Touren geht von den geographischen Koordinaten der Auslieferstellen aus. Eine von Ahuja vorgeschlagene einfache Methode zur Clusterung von Punktmengen in der euklidischen Ebene ist die Zerlegung des *minimal aufspannenden Baumes* einer Punktmenge in einen aufspannenden Wald. Von diesem Ansatz ausgehend wird für eine vorgegebene Instanz des Vehicle-Routing-Problems ein Baumzerlegungsproblem unter Nebenbedingungen formuliert. Dazu berechnet man zunächst einen minimal aufspannenden Baum für den vollständigen Graphen, dessen Knoten den Kunden entsprechen und dessen Kantengewichte durch die Fahrzeiten im Straßennetz zwischen je zwei Kunden gegeben sind.

Dieser Baum wird nun mit Hilfe eines exakten Verfahrens auf der Basis eines dynamischen Programmes in gültige Teilbäume zerlegt. Jeder der Teilbäume repräsentiert eine Gruppe von Kunden, die mit dem gleichen Fahrzeug beliefert werden soll. Die Bedingungen an die Touren werden mit Hilfe der Knotengewichte formuliert. Für die Qualität der Clusterung ist die problemadäquate Modellierung der Gewichtsfunktionen unerläßlich.

Ergebnisse

Die Realisierung des dynamischen Programms zur exakten Lösung des Baumzerlegungsproblems zeigte sowohl auf zufälligen als auch auf den Probleminstanzen aus der Praxis hervorragendes Laufzeitverhalten. Die theoretischen Worst-Case-Schranken wurden deutlich unterschritten. Insbesondere die Praxisinstanzen von etwa 1000 Kunden mit fünf Knotengewichtsfunktionen ließen sich binnen weniger Sekunden exakt lösen. Uns ist es bisher nur gelungen, für den Fall zweier Gewichtsfunktionen auf binären Bäumen ein Worst-Case-Beispiel zu konstruieren. Die allgemeinen theoretischen Schranken scheinen noch Raum für eine strengere Laufzeitanalyse zu lassen.

1.3.3 Zusammenarbeit mit der Firma Profi.S

Die Baumzerlegungsheuristik wurde von der Tourenplanungsgruppe in das Programm „ProfiTour“ der Firma Profi.S integriert. Dieses Tourenplanungswerkzeug wird in verschiedenen Speditionen, insbesondere der Möbelindustrie, eingesetzt. Für dieses Programm wurde im Berichtszeitraum vom ZAIK/ZPR eine Fahrzeugeinsatzplanung mitentwickelt, die Anforderungen wie den Wiedereinsatz von LKW oder inhomogene Fahrzzeugflotten in den Optimierungsalgorithmen berücksichtigt. Da im Möbelbereich in vielen Fällen direkt vom Band in den LKW produziert wird, wurden außerdem erste Schritte zur algorithmischen Verknüpfung von Produktions– und Tourenplanung implementiert. Die Zusammenarbeit mit der Profi.S garantiert die Anwendernähe der am ZAIK/ZPR entwickelten innovativen Technologien für den Tourenplanungsbereich.

1.4 Rendezvous-Suche

In vielen alltäglichen Situationen ist man mit dem Problem konfrontiert, eine andere Person in einer unübersichtlichen Umgebung zu treffen. Genaue Verabredungen helfen, Schwierigkeiten zu verkleinern, aber oft kann man keinen präzisen Ort und keine präzise Zeit vereinbaren, da man den eigenen späteren Aufenthaltsort zum Zeitpunkt der Verabredung noch nicht kennt. Szenarien dieser Art treten aber auch bei Rettungsaktionen auf, bei denen zumindest einer der Beteiligten nicht genau weiß, wo der andere sich gerade befindet. Probleme mit zwei oder mehr mobilen Personen („Spielern“) bezeichnet man als Rendezvous-Suche. Ziel ist es, optimale Protokolle zu entwickeln, die ein möglichst schnelles Zusammentreffen garantieren, bzw. die Wahrscheinlichkeit für ein Zusammentreffen maximieren.

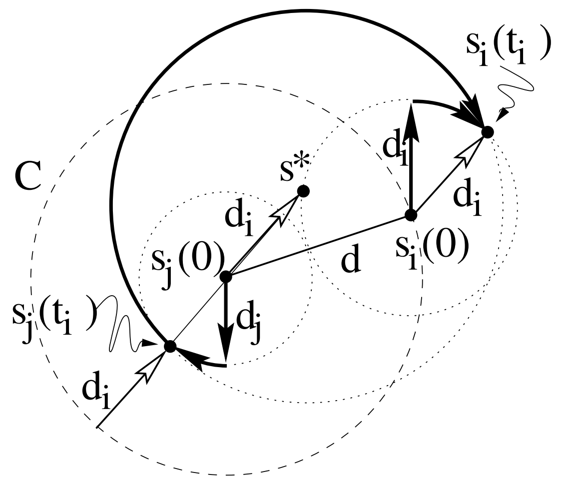

Abbildung 1.6: Eine optimales Strategiepaar für eine begrenzte Gesamtstrecke

In vergangenen Jahren gab es eine zunehmende Anzahl von Untersuchungen zur Rendezvous-Suche in eindimensionalen Szenarien, bei denen sich die beiden Spieler z. B. auf einer Strecke oder einem Kreis befinden. In Zusammenarbeit mit Prof. Edward Anderson (Sydney) gelang es uns, erstmals eine Reihe von Ergebnissen für zweidimensionale Rendezvous Probleme zu erzielen. Unterschiedliche Szenarien ergeben sich durch verschiedene Annahmen über die den Spielern zur Verfügung stehende Information. So kann der anfängliche Abstand bekannt sein, oder es kann eine gemeinsame Orientierung im Raum existieren. Für diese Szenarien gelang es uns, eine optimale Strategie zu finden, die sowohl den Erwartungswert bei zufälligem Ausgangspunkt minimiert, als auch bei begrenzter Gesamtstrecke beider Beteiligten die Wahrscheinlichkeit für ein Treffen maximiert. Für den Fall, daß sich die Spieler auf einem Gitter bewegen, fanden wir ebenfalls eine optimale Strategie, auch wenn die Anfangsorientierung zufällig ist. Schließlich gelang es uns, ein typisches Rettungsszenario zu lösen, bei dem einer der beiden Spieler (der „Retter“) die Anfangsposition und den Abstand des anderen Spielers (des „Verirrten“) kennt, während dieser nur die Anfangsdistanz des Retters kennt.

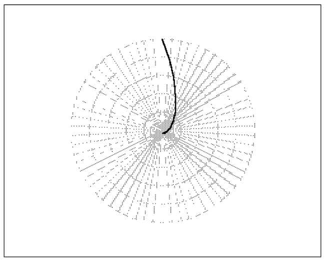

Abbildung 1.7: Eine optimale Trajektorie für das Rettungsszenario

1.5 Computational Finance

Im modernen Bankbereich und der Firmenwelt geben die immer komplizierter werdenden Produkte auf dem Markt der Derivate einen starken Anreiz für neue mathematische Modelle und Algorithmen. In der Computational Finance werden Probleme und Modelle aus dem Blickwinkel des

Mathematikers/Informatikers untersucht, wobei die diskrete Modellierung und algorithmische Fragen zur Lösung der Probleme im Vordergrund stehen.

In Zusammenarbeit mit Ming Yang Kao (Department of Computer Science, Yale University) studieren wir das Problem, eine optimale Investment Strategie für Investoren mit unterschiedlichen Einschätzungen des Marktes zu finden. Die Einzelverteilungsfunktionen der Erträge der Aktien, die von den Investoren betrachtet werden, sind bekannt, während die gemeinsame Verteilung unbekannt ist. Das Problem ist, die beste Investment Strategie zu finden, so daß die Wahrscheinlichkeit, einen gewissen Prozentsatz des eingesetzten Kapitals zu verlieren, minimiert und gleichzeitig der Einstellung der Investoren gegenüber zukünftigen Entwicklungen des Marktes Rechnung getragen wird.

Für besonders pessimistische oder besonders optimistische Investoren läßt sich das Problem im zweidimensionalen Fall als Flußproblem modellieren und erstaunlicherweise durch einen einfachen Greedy-Ansatz lösen. Für höherdimensionale Fälle läßt sich das Problem nicht mehr in polynomieller Zeit bearbeiten, jedoch kann man noch einen exakten Algorithmus angeben, der polynomiell in der Anzahl der Einträge der Matrix einer gemeinsamen Verteilung ist. Als Konsequenz entwickelten wir einen Approximationsalgorithmus, der das Problem nur annähernd, jedoch in polynomieller Zeit, löst. Falls der Investor indifferent gegenüber der zukünftigen Marktentwicklung ist, läßt sich das Problem im zweidimensionalen Fall durch einen auf einem random walk basierenden Algorithmus lösen, es wird allerdings auch hier für höhere Dimensionen NP-schwer.

1.5.1 Orthogonale Packungsprobleme

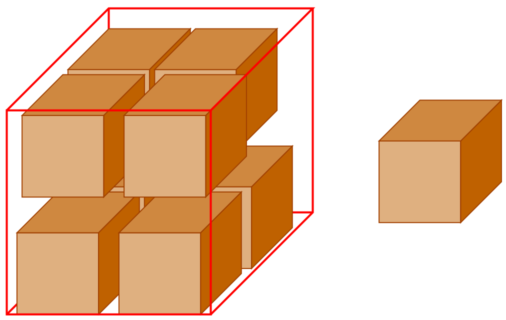

Abbildung 1.8: Lassen sich 9 Quader der Seitenlänge 0.4 in einen Einheitswürfel packen?

Zahlreiche in Industrie und Wirtschaft auftretende Fragestellungen lassen sich als mehrdimensionale Packungsprobleme formulieren. Neben Aufgaben wie dem Beladen von Paletten und Containern, dem Erstellen von Platinenlayouts oder dem Zuschnitt von Materialien (Cutting Stock Probleme) zählen hierzu auch Schedulingprobleme mit partitionierbaren Ressourcen. Die meisten dieser Fragestellungen fallen in die Kategorie der orthogonalen Packungsprobleme. Dabei versucht man, eine Menge d-dimensionaler massiver Quader in einem oder mehreren quaderförmigen Behältern achsenparallel „möglichst günstig“ zu plazieren. So will man beim orthogonalen Bin-Packing-Problem möglichst wenige Behälter verwenden, beim orthogonalen Knapsackproblem den Gesamtwert der Güter in einem Behälter maximieren oder beim Strip-Packing-Problem die Höhe der Packung minimieren.

Diese Probleme sind als Verallgemeinerungen des eindimensionalen Bin-Packing-Problems NP-schwer im strengen Sinne. Dennoch vermittelt die komplexitätstheoretische Einordnung allein noch keinen Eindruck von ihrem immensen Schwierigkeitsgrad. Während man gegenwärtig für das im weiteren Sinne NP-schwere eindimensionale Knapsackproblem Instanzen mit bis zu 250.000 Objekten exakt lösen kann, umfaßte die größte bisher exakt gelöste Instanz des zweidimensionalen orthogonalen Knapsackproblems ganze 22 Objekte.

Zur Lösung praktischer Probleme werden daher heuristische Methoden eingesetzt. In der Regel produzieren diese Verfahren zufriedenstellende Lösungen. Entscheidend ist dabei allerdings, daß die zu packenden Gegenstände wesentlich kleiner als die Behälter sind. Beim Packen weniger sperriger Teile kann bereits ein einziges ungünstig plaziertes Objekt die Qualität der Gesamtlösung drastisch verschlechtern. Daher ist es erforderlich, Instanzen mit wenigen sperrigen Teilen exakt zu lösen.

1.5.2 Modellierung

Die in der Literatur beschriebenen exakten Verfahren beruhen auf ganzzahliger linearer Programmierung. Sie sind nur für zweidimensionale Probleme verwendbar und stoßen bereits bei sehr kleinen Instanzen an ihre Grenzen. In unseren Untersuchungen entwickelten wir eine neue Art der Modellierung zulässiger Packungen. Es gelang uns zu zeigen, daß sich jede zulässige Anordnung durch einige einfache graphentheoretische Bedingungen charakterisieren läßt.

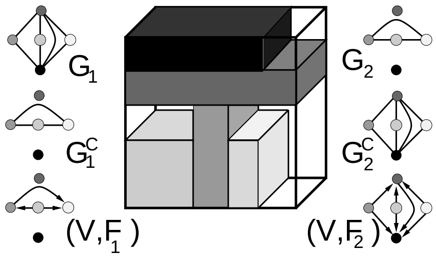

Abbildung 1.9: Eine zulässige Packung und die zugehörigen Graphen

Damit kann man die aus der Geometrie entstehenden Schwierigkeiten deutlich reduzieren und sich statt dessen auf die Behandlung der aus der Charakterisierung resultierenden Graphen konzentrieren. Ein Vorzug ist, daß zu einer Auswahl von Graphen normalerweise eine sehr große Zahl von zulässigen Packungen gehört, d. h. man nutzt bei einer Suche mittels Branch-&-Bound allein durch die Wahl dieser Datenstruktur bereits ein großes Maß an Symmetrien aus. Außerdem kommt einem die Tatsache zugute, daß die zu untersuchenden Graphen einige aus der klassischen Kombinatorik bekannte, algorithmisch besonders gutartige Eigenschaften haben.

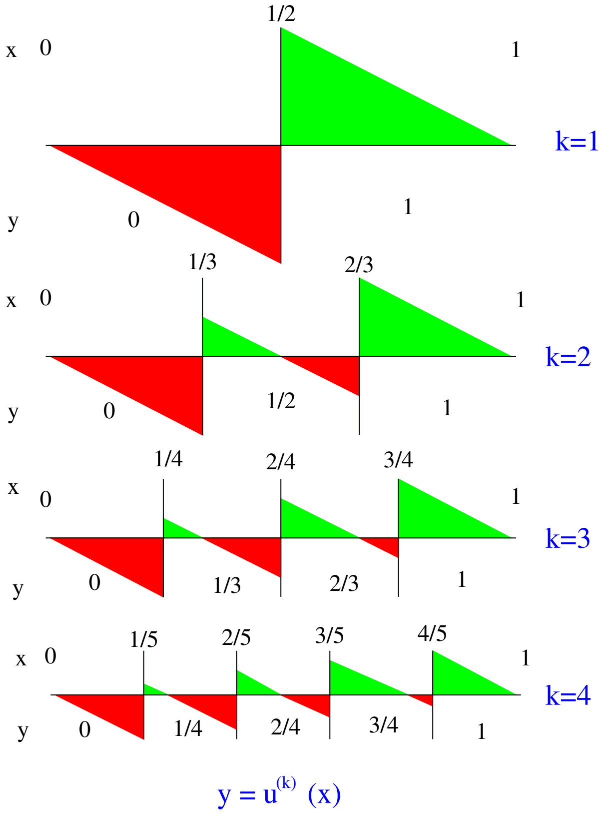

Abbildung 1.10: Die Volumenbilanz für eine Familie dualzulässiger Funktionen

1.5.3 Schranken

Für einen erfolgreichen Branch-&-Bound-Algorithmus sind gute Schranken, die sich leicht berechnen lassen, von besonderer Bedeutung. Sie erlauben es, die Baumsuche effizient durchzuführen, da sich große Teile des Suchbaumes schnell als irrelevant für die Suche nach einem Optimum identifizieren lassen. Neben einfachen Schranken für den eindimensionalen Fall gab es bislang eine Reihe von „handgestrickten“ Varianten für zwei- oder dreidimensionale Probleme. Die Gültigkeit dieser Schranken mußte bislang aber jeweils einzeln und relativ mühsam bewiesen werden.

Durch Untersuchungen am klassischen eindimensionalen Bin-Packing-Problem gelangten wir zu einem neuen allgemeinen Ansatz, mit dem sich die Entwicklung unterer Schranken systematisieren und vereinfachen läßt. Gleichzeitig erhielten wir mit unserer Methode der sogenannten „dualzulässigen Funktionen“ eine Reihe von verallgemeinerten Schranken, die sowohl theoretisch wie praktisch eine Reihe von nützlichen Eigenschaften aufweisen: Sie sind extrem schnell und einfach zu berechnen, sie liefern das Optimum für ganze Klassen von Probleminstanzen, für die andere Typen von Schranken besonders schlecht abschneiden, sie besitzen auch im schlechtesten Falle eine hervorragende Gütegarantie und schließlich schneiden sie auch in etablierten Praxistests deutlich besser ab als bisher gebräuchliche einfache Schranken.

1.5.4 Exakte Algorithmen

Die Nützlichkeit der verschiedenen theoretischen Neuerungen auch im Zusammenspiel konnten wir durch eine Implementation des resultierenden Gesamtverfahrens zeigen. Für das zweidimensionale orthogonale Knapsackproblem konnte die maximale Anzahl der Objekte in einer exakt gelösten Instanz von 22 auf 80, und die maximale Anzahl verschiedener Objektgrößen von 15 auf 40 gesteigert werden. Diese Zahlen geben keine singulären Extremfälle wieder, sondern beziehen sich auf ganze Klassen nichttrivialer Testinstanzen. Nach unserem Wissen haben wir erstmalig experimentelle Resultate für das dreidimensionale

orthogonale Knapsackproblem veröffentlicht.
Bei einem geringen Anteil kleiner Quader konnten hier ebenso große Instanzen bearbeitet werden wie im zweidimensionalen Fall.

1.5.5 Optimierung dynamischer Hardware-Rekonfiguration

Inzwischen läuft ein Projekt zur weiteren Anwendung unserer Packalgorithmen in der Praxis. Gemeinsam mit Prof. Jürgen Teich vom Fachbereich Elektrotechnik der Universität-Gesamthochschule Paderborn arbeiten wir am Einsatz unserer Packverfahren zur optimalen Belegungsplanung besonderer Computerchips, sogenannter „Field Programmable Gate Arrays“ (FPGAs). Diese Bauteile ermöglichen es, zahlreiche Berechnungen parallel laufen zu lassen. Nach Abschluß eines einzelnen Jobs läßt sich der dadurch freiwerdende Teilbereich des Chips softwaregesteuert für weitere Berechnungen neu konfigurieren, ohne daß andere laufende Berechnungen gestoppt werden müssen. Probleme dieser Art lassen sich als dreidimensionale Packungsprobleme mit Nebenbedingungen formulieren, wobei zwei Dimensionen die Geometrie des Chips repräsentieren (ein einzelner Job nimmt also einen rechteckigen Teilbereich ein), während die dritte Dimension für die Zeit steht.

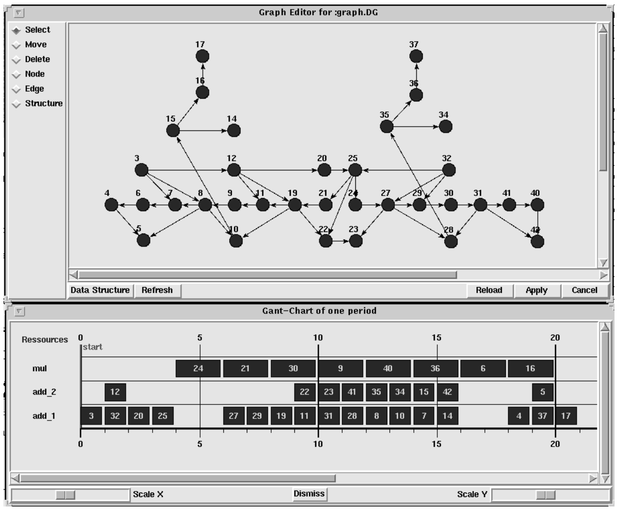

Abbildung 1.11: Restriktionsgraph und optimaler Belegungsplan für die Berechnungen eines elliptischen Wellenfilters

Aufgrund unserer graphentheoretischen Charakterisierung sind die aus Reihenfolgerestriktionen für einzelne Berechnungen resultierenden Nebenbedingungen mit unserem Verfahren relativ gut zu verarbeiten, da sie auf natürliche Weise den Suchbaum einschränken. Statt eines dreidimensionalen Problemes mit schwer zu erfüllenden Zusatzbedingungen können wir das Problem in gewissen Szenarien auf ein zweidimensionales Problem reduzieren. Erste Versuche mit etablierten Benchmarkproblemen ergaben ermutigende Resultate.

1.6 Geometrische Probleme

1.6.1 Kontinuierliche Standortprobleme in Polygonen

Das mathematische Problem der planaren Standortplanung taucht in einer Reihe von praktischen Zusammenhängen auf, z. B. in der Ökonomie. Dabei geht es darum, aus einer Menge von möglichen Standorten so eine Auswahl zu treffen, daß der Abstand zu einer die Nachfrage beschreibenden Menge optimiert wird. Die zahlreichen bisher in der Literatur veröffentlichten Untersuchungen gehen von diskreten Nachfrageverteilungen aus. Wir betrachten nun kontinuierliche Nachfrageverteilungen: als Zielfunktion wird der durchschnittliche Abstand für alle Punkte innerhalb eines betrachteten Polygons verwendet, diese stellen gleichzeitig die Menge zulässiger Standorte dar.

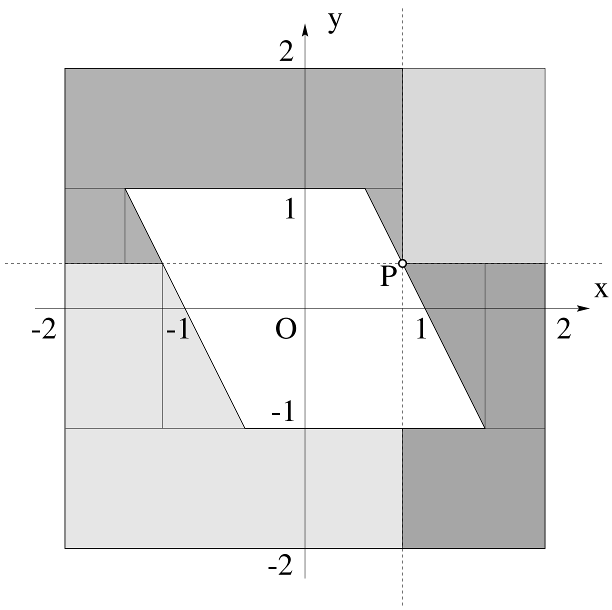

Abbildung 1.12: $P(x, y)$ und sein Spiegelpunkt $P'(-x, -y)$ sind die beiden Minimalstellen für den L_1–Abstand

Wir erhielten sehr unterschiedliche Ergebnisse für unterschiedliche Fragestellungen: Wird der Abstand mit der sogenannten Manhattan-Metrik gemessen, so läßt sich für ein zusammenhängendes Plazierungspolygon ohne Löcher die Plazierung eines einzigen Zentrums mit möglichst kleinem durchschnittlichen Abstand durch einen sehr schnellen Algorithmus lösen. Im Falle von Bereichen mit Löchern läßt sich immerhin noch ein Algorithmus angeben, der die Aufgabe in $O(n^2)$ löst; dieser läßt sich auch anwenden, wenn der Plazierungsbereich nicht zusammenhängend ist. Wird der Abstand anhand kürzester Wege gemessen, die den zulässigen Bereich nicht verlassen dürfen (sogenannte *geodätische Abstände*), so lassen sich die entsprechenden Fragestellungen für einen zusammenhängenden Bereich mit jeweils wenig mehr Aufwand beantworten. Sollen mehrere Zentren gleichzeitig plaziert werden, wird die Aufgabe gleich sehr viel komplexer: Für geodätischen Manhattan–Abstand und Polygone mit Löchern können wir zeigen, daß das Problem NP–schwer ist. Für euklidische Abstände sind bereits einfachste Fragestellungen mit lediglich fünf diskreten Nachfragepunkten und uneingeschränkter Plazierungsmenge in einem bestimmten algebraischen Sinne nicht algorithmisch lösbar, so daß man für die Verallgemeinerung auf kontinuierliche Nachfrageverteilung nicht auf effiziente algorithmische Lösungen hoffen kann.

1.6.2 Sterne und Matchings

Wenn man für eine gegebene endliche Menge von Punkten ein optimales Zentrum plaziert, das die Summe der Abstände minimiert, so kann man die Verbindungen vom Zentrum zu den Punkten auch als einen speziellen Baum auffassen, einen sogenannten *Stern*. Ist das Zentrum nicht notwendigerweise ein Punkt der gegebenen Menge, so sprechen wir (in Anlehnung an sogenannte *Steinerbäume*) von einem *Steinerstern*. Typischerweise ist man an Sternen und Steinersternen möglichst kleiner Gesamtlänge interessiert.

Ein anderes klassisches Optimierungsproblem für eine Menge von Punkten ist das sogenannte *Matching*. Hierbei bildet man eine Menge von disjunkten Paaren und addiert die Distanzen der Punkte aller Paare. Es ist nun eine einfache Folge der Dreiecksungleichung, daß die Länge $\min ||StSt||$ eines minimalen Steinersternes nicht kürzer sein kann als die Länge $\max ||Mat||$ eines maximalen Matchings. Für den Fall von Manhattan-Distanzen in der Ebene gilt sogar Gleichheit, für den Fall von euklidischen Distanzen sieht man leicht, daß es Beispiele gibt, für die $\min ||StSt|| / \max ||Mat|| > 1$ gilt.

Motiviert durch Probleme aus dem Kontext optimaler Kommunikationsnetzwerke hatte Suri bereits 1995 nach dem größtmöglichen Wert dieses Verhältnisses gefragt und wiederholte diese Frage im Juni 1998 als besondere Herausforderung auf der Jahrestagung für algorithmische Geometrie. In Zusammenarbeit mit Prof. Henk Meijer (Queen's University, Kanada) gelang es uns zu beweisen, daß das genannte Verhältnis den Wert $2/\sqrt{3} \approx 1.17$ nicht überschreiten kann, was die bestmögliche Abschätzung ist.

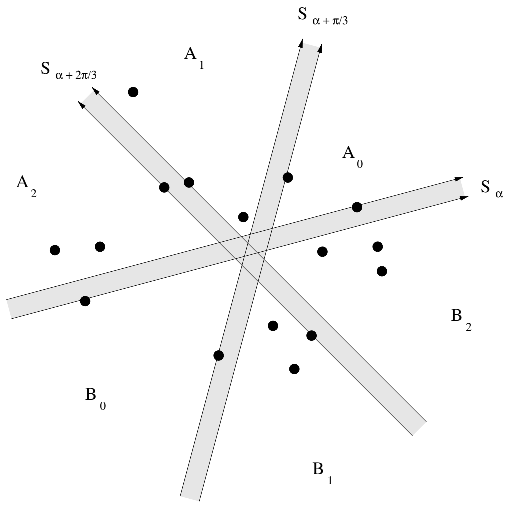

Abbildung 1.13: Finden eines kleinen Sternes und eines großen Matchings

Darüber hinaus fanden wir eine ganze Reihe von weiteren (zum Teil scharfen) Abschätzungen für die Verhältnisse zwischen minimalen Sternen, minimalen Steinersternen und maximalen Matchings; je nach Metrik und Dimension des Raumes ergeben sich dabei sehr unterschiedliche Werte.

1.6.3 Dispersionsprobleme

Während es bei Packungsproblemen darum geht, möglichst viele Objekte auf möglichst kleinem Raum unterzubringen, ist bei Dispersionproblemen das Ziel, die Objekte möglichst weit voneinander entfernt zu plazieren. Fragestellungen dieser Art treten auf, wenn sich die Objekte so wenig wie möglich gegenseitig stören sollen, etwa bei Rundfunksendern oder Zweigstellen einer Ladenkette. Dispersionsprobleme lassen sich auch als Packungsprobleme modellieren, bei denen man eine fixe Anzahl von möglichst großen Objekten zu plazieren hat. Vom mathematischen Standpunkt sind fast alle dieser Probleme NP-schwer, d. h. es ist nicht damit zu rechnen, daß es einen exakten Algorithmus mit schneller Laufzeit gibt.

Hochbaum und Maass stellten 1985 ein polynomielles Approximationsschema vor, das für einen bestimmten Typ von Packungsproblemen schnelle Approximationsalgorithmen mit Gütegarantie beliebig nahe bei 1 liefert. Demgegenüber war für Dispersionsprobleme ein derartiges Schema nicht bekannt und es war unklar, ob es solch ein Schema überhaupt geben könne. Es gelang uns, diese offene Frage zu lösen, indem wir zeigten, daß es unter der Voraussetzung $P \neq NP$ keinen polynomiellen Approximationsalgorithmus mit Gütegarantie kleiner als 14/13 geben kann. Für eine besondere Klasse von Problemen liegt diese Schranke sogar bei 2, was sich mit der Gütegarantie des besten bekannten Approximationsalgorithmus deckt. Dies löste ein offenes Problem von Ravi, Rosencrantz und Tayi. Andererseits gelang es uns für eine andere Problemklasse, diesen besten bekannten Approximationsfaktor zu verbessern, indem wir einen polynomiellen Approximationsalgorithmus mit Gütegarantie 3/2 entwickelten.

1.6.4 Spanner in planaren Graphen

Bei der Optimierung eines Netzwerkes tritt oft die Situation ein, daß man die Menge der bestehenden direkten Verbindungen reduzieren muß, um Kosten zu sparen. Dabei ist es meist unerwünscht, daß sich die Länge des kürzesten Weges zwischen zwei Knoten des Netzwerkes gegenüber der ursprünglichen Entfernung um einen zu großen Faktor erhöht. Für ein Teilnetzwerk bezeichnet man den größten derartig auftretenden Faktor als den sogenannten *Streckungsfaktor*; sind alle ursprünglich voneinander erreichbaren Knoten auch weiterhin verbunden (wenn auch durch möglicherweise durch längere Wege), so nennt man das Teilnetzwerk einen *Spanner* des Netzwerkes. Spanner werden schon seit einer Reihe von Jahren algorithmisch untersucht. Neben der Bedeutung für die Netzwerkoptimierung treten sie auch überall da auf, wo man eine große Menge von Informationen über Einzelverbindungen durch eine wesentlich kleinere Zahl ersetzen will, ohne daß dadurch die Information über Kosten oder Entfernungen zu stark geändert wird.

Ist man an Spannern minimalen Gewichts interessiert, so sind Baumspanner besonders interessant, d. h. Teilnetzwerke, aus denen keine weiteren Kanten mehr entfernt werden können, ohne daß ursprünglich vorhandene Verbindungen vollständig verloren gehen. Cai und Corneil haben vor einigen Jahren gezeigt, daß die Existenz eines 2-Baumspanners in beliebigen Netzwerken in polynomieller Zeit überprüft werden kann, die Entscheidung über die Existenz eines 4-Baumspanners hingegen NP-vollständig ist. Die Komplexität von 3-Baumspannern in beliebigen Netzwerken ist schon seit längerem offen.

Eine Klasse von Netzwerken von großer Bedeutung sind die sogenannten *planaren Graphen* – also Netzwerke, die ohne Überkreuzung von Kanten in der Ebene darstellbar sind. Da es in planaren Graphen besonders einfache Beziehungen zwischen den in einem Baumspanner verbleibenden und den entfernten Kanten gibt (die entfernten Kanten bilden einen Baumspanner des Dualgraphen), war es durchaus denkbar, daß für diese Graphenklasse schnelle Algorithmen zur Auffindung guter Baumspanner existieren. Es gelang uns aber nachzuweisen, daß die Frage nach einem Baumspanner mit optimalem Stretchfaktor bereits für planare Graphen NP-schwer ist. Andererseits entwickelten wir einen Algorithmus, der in polynomieller Zeit entscheidet, ob ein planarer Graph einen 3-Baumspanner hat. Zumindest für diesen Spezialfall konnten wir also eine positive Antwort auf das Problem von Cai und Corneil finden.

1.6.5 Das Maximum-Traveling-Salesman-Problem

Das sogenannte *Traveling-Salesman-Problem* (TSP, auch Problem des Handlungsreisenden genannt) ist wahrscheinlich das bekannteste Problem der Kombinatorischen Optimierung. Es ist leicht zu beschreiben: „Plane eine möglichst kurze Rundreise durch eine gegebene Menge von Städten.“ Interessant wird das TSP vor allem dadurch, daß das Problem trotz dieser einfachen Beschreibung nur sehr schwer zu lösen ist, da es zur Klasse der sogenannten NP-schweren Probleme gehört. Wissenschaftliche Ergebnisse am TSP haben schon mehrfach zu Durchbrüchen bei theoretischer wie praktischer Forschung in der Optimierung geführt. So konnte erst vor kurzem von Arora und Mitchell gezeigt werden, daß sich für geometrische TSP-Instanzen (in denen die Städte Punkten im Raum mit ihren geometrischen Distanzen entsprechen) sogenannte polynomielle Approximationsschemata existieren – d. h., für jede beliebige Genauigkeit gibt es in einem bestimmten Sinne schnelle Algorithmen, die eine Optimallösung gut approximieren.

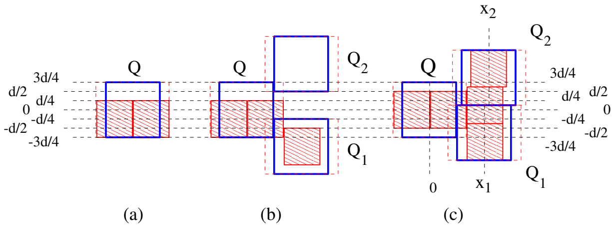

Abbildung 1.14: Konstruktion des 2/3-Approximationsalgorithmus

Vom mathematischen Standpunkt erscheint das sogenannte *Maximum-TSP* eng verwandt: Hier geht es darum, eine längste Rundreise zu finden. Tatsächlich sind die betrachteten graphentheoretischen Objekte (sogenannte *Hamiltonkreise*) dieselben, und es ist leicht zu zeigen, daß das Maximum-TSP im allgemeinen ebenfalls NP-schwer ist. Nachdem Barvinok vor einigen Jahren für geometrische Instanzen ein polynomielles Approximationsschema entwickeln konnte, war es eine echte Sensation, als ihm im Jahre 1998 zusammen mit Johnson, Woeginger und Woodroofe der Nachweis gelang, daß bei bestimmten gängigen Abstandsmaßen in Räumen konstanter Dimension polynomielle Algorithmen existieren. Dazu gehören z. B. die sogenannten Manhattan-Metriken, bei denen die Abstände für Wege parallel zu den Koordinatenachsen gemessen werden. Dagegen war die Komplexität des Maximum-TSP für die gewöhnliche euklidische Metrik im zwei- oder mehrdimensionalen Raum ein offenes Problem.

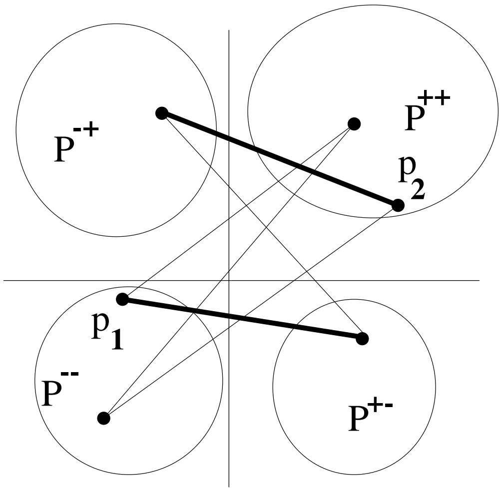

Abbildung 1.15: Eine optimale Tour, die sich schnell finden läßt

Zunächst gelang uns durch einen Beweis der NP-Schwere für das Euklidische Maximum-TSP im dreidimensionalen Raum die Lösung des Problems von Barvinok et al. und damit der Nachweis, daß selbst für ganz natürliche geometrische Optimierungsprobleme die algorithmische Schwierigkeit ausschließlich von der Geometrie des Abstandsmaßes abhängen kann statt nur von der Kombinatorik des Problems.

Für den Fall von Manhattan-Distanzen in der Ebene entwickelten wir dann sogar noch einen einfachen Algorithmus mit linearer, also schnellstmöglicher Laufzeit. Dieses überraschende Resultat unterstreicht die wichtige Rolle der Geometrie in der Behandlung kombinatorischer Optimierungsprobleme.

1.7 Randomisierte Algorithmen

Im letzten Jahrzehnt haben randomisierte Algorithmen enorm an Bedeutung gewonnen. Während dieser Periode sind sie von einem Werkzeug aus der Zahlentheorie zu einem der bedeutensten algorithmischen Konzepte mit weitreichenden Anwendungen in verschiedenen Gebieten geworden. Randomisierte Algorithmen sind dabei Algorithmen, die sich im Gegensatz zu deterministischen Algorithmen den Zufall zunutze machen. Zwei Vorteile gehen in der Regel mit randomisierten Algorithmen einher, Einfachheit und Geschwindigkeit. Für viele Anwendungen sind die randomisierten Algorithmen entweder die einfachsten, die schnellsten oder vereinen beide Vorteile.

Unsere Forschung, in Kooperation mit Ravi Kannan (Department of Computer Science, Yale University) und Leen Stougie (Department of Computer Science, Eindhoven University) konzentrierte sich auf drei Anwendungen von randomisierten Algorithmen:

Das Component Commonality Problem

Das *Component Commonality Problem* ist ein stochastisches Optimierungsproblem aus dem Bereich supply chain optimization, das wie folgt beschrieben werden kann. Es gibt m Produkte, deren zugeordneter Bedarf dem Zufall unterliegt. Jedes dieser Produkte benötigt zur Herstellung einige oder alle der n Komponenten in einer Periode. Die Schwierigkeit liegt darin, daß die Komponenten zu Beginn einer Periode gekauft werden müssen, wenn noch keine Information über den tatsächlichen Bedarf vorhanden ist. Der einzige Hinweis auf den zukünftigen Bedarf ist die Information über die Verteilung des Bedarfs in der Vergangenheit. So könnten zum Beispiel Bedarfsvektoren aus vielen Perioden der Vergangenheit gesammelt worden sein. Das Problem besteht nun darin, die Komponenten so einzukaufen, daß die Kosten minimiert werden, jedoch gleichzeitig die Wahrscheinlichkeit, daß der Bedarf befriedigt wird, nicht unter eine gewisse vorgegebene Grenze sinken darf.

Es wurde ein Algorithmus entwickelt, der dieses Problem approximativ löst. Der Algorithmus basiert auf einem geometrischen random walk und stellt den Algorithmus mit der besten zur Zeit bekannten Zeitkomplexität für dieses Problem dar.

Smooth Sets

Auf diesem Gebiet wurden zwei sehr einfache Techniken entwickelt, eine lineare Funktion über einer konvexen Menge zu optimieren. Die erste ist ein deterministischer Algorithmus, der auf der Gradientenmethode basiert. Die zweite ist ein randomisierter Algorithmus, der kleine lokale und zufällige Änderungen in jedem Schritt durchführt. Die zweite Methode kann benutzt werden, wenn die Menge nur durch ein Membership-Orakel beschrieben ist, während die erste Methode etwas ähnliches wie ein Separationsorakel erfordert. Wir entwickelten eine geeignete Definition von smooth sets und waren in der Lage nachzuweisen, daß beide Algorithmen nahezu optimale Lösungen in polynomieller Zeit liefern. Diese Optimierungsmethode hat weitreichende Anwendungen in der Linearen Optimierung und Stochastischen Optimierung, wo die relevanten Mengen in der Tat smooth und daher unsere Algorithmen anwendbar sind. Trotz ihrer simplen Beschreibung liefern sie in vielen Fällen die beste bisher bekannte Zeitkomplexität.

Clustering

Das *Clustering Problem* besteht darin, m Punkte im n-dimensionalen Raum jeweils einem von k Zentralpunkten so zuzuordnen, so daß die Summe der Abstände der Punkte von ihren zugeordneten Zentralpunkten minimiert wird. Dieses Problem hat eine Fülle von Anwendungen wie zum Beispiel in der Schrifterkennung, bei der einer Menge von handgeschriebenen Buchstaben die korrekten Buchstaben zugeordnet werden müssen. Aber auch in der Versicherungsbranche finden sich Anwendungen, wenn man die Menge der Versicherten in Gruppen ähnlichen Schadensverhaltens aufteilen will, um so die Prämien gerechter gestalten zu können.

Auf diesem Gebiet wurde ein auf einem Sampling Ansatz basierender Algorithmus entwickelt, der für ein bestimmtes Abstandsmaß ein polynomielles Approximationsschema liefert und somit auf diesem Gebiet der erste Algorithmus überhaupt ist, der in akzeptabler Zeit eine Lösung mit garantierter Güte liefert.

1.8 Testmengen in der ganzzahligen Optimierung

Die Kenntnis von Testmengen ermöglicht, ganzzahlige Optimierungsprobleme mittels eines primalen Optimierungsalgorithmus zu lösen. Testmengen sind durch folgende Eigenschaft charakterisiert: Eine zulässige Lösung x eines ganzzahligen Optimierungsproblems ist entweder optimal, oder es gibt ein Element t der Testmenge, so daß die Differenz x-t wieder zulässig ist und bezüglich der Zielfunktion einen besseren Wert hat. Ist also eine zulässige Lösung gegeben, kann man iterativ durch Subtraktion von Testmengenelementen zur Optimallösung gelangen.

Conti und Traverso haben einen interessanten Zusammenhang zwischen ganzzahliger Optimierung und Gröbnerbasen in der kommutativen Algebra aufgezeigt. Dabei werden die Differenzen von zulässigen Lösungen zu einer ganzen Familie von geeigneten ganzzahligen Programmen in Beziehung zu einem Binomideal in einem Polynomring gesetzt. Bei geeigneter Wahl einer Monomordnung entspricht dann die reduzierte Gröbner-Basis dieses Ideals der minimalen Testmenge für die Problemfamilie. Theoretisch ergibt sich dadurch die Möglichkeit, mit Hilfe algebraischer Methoden ganzzahlige Optimierungsprobleme zu lösen: Erst wird mit dem Buchberger-Algorithmus die Testmenge ermittelt. Anschließend wird ausgehend von einer zulässigen Lösung eine Folge von Verbesserungsschritten „entlang“ Elementen der Testmenge ausgeführt, bis das Optimum erreicht ist.

Obwohl in dem betrachteten Fall die Gröbnerbasis nur zu Idealen mit recht spezieller Struktur berechnet werden muß, ist dieser Ansatz aufgrund der hohen Laufzeit

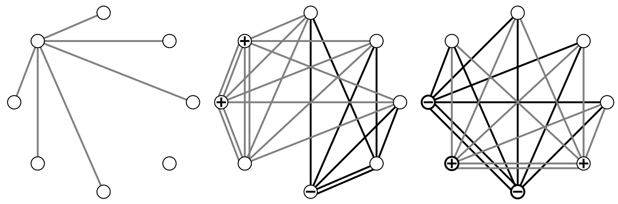

Abbildung 1.16: Repräsentanten der drei Klassen von Elementen in der minimalen Testmenge für K_8

des Buchberger Algorithmus und der potentiell gewaltigen Größe der Basis für Beispiele mit vielen Variablen sehr problematisch. Daher werden die algebraischen Methoden nicht als direkte Lösungsverfahren betrachtet. Statt dessen werden die Methoden der Computeralgebra im hier verfolgten Ansatz dazu eingesetzt, die minimalen Testmengen zu ganzzahligen Optimierungsproblemen zu berechnen, um anschließend aus den so gewonnenen Informationen Aussagen über die Struktur der Testmengen zu ermöglichen. So ist es beispielsweise im Falle der Bestimmung einer Knotenüberdeckung minimalen Gewichts in einem Graphen gelungen, die Struktur der Testmenge vollständig zu beschreiben, und zwar unabhängig von der Tatsache, daß zu ihrer Berechnung aufwendige algebraische Methoden eingesetzt wurden. Für den Spezialfall einer minimalen Knotenüberdeckung in einem vollständigen Graphen gilt der Satz:

Bis auf ein Element ist die reduzierte Gröbnerbasis genau die Menge aller Vektoren, die entweder negativ bzgl. der Monomordnung sind und das Gewicht um 1 erniedrigen, oder das Gewicht unverändert lassen und positiv bzgl. der Ordnung sind.

In Abbildung 1.8 sind typische Repräsentanten der minimalen Testmenge für den K_8 graphisch dargestellt. Das linke Element und Elemente vom mittleren Typ erniedrigen das Gewicht einer Knotenüberdeckung, während die Elemente vom rechten Typ lediglich für eine Umverteilung der Gewichte stehen. Im Falle beliebiger Graphen kommt noch eine dritte Klasse von Elementen zur Testmenge hinzu, die die komplizierteren Adjazenz-Beziehungen zwischen Knoten mit veränderten Gewichten widerspiegelt.

Ziel ist es, auch bei schweren ganzzahligen Optimierungsproblemen eine – möglicherweise unvollständige – Beschreibung der Struktur der Testmenge ermitteln zu können. Diese Kenntnisse können dann in einer Verbesserungsheuristik genutzt werden.

1.9 Binäre Matroide

Zu jedem Graphen kann man kanonisch eine Punktkonfiguration im $|V|$-dimensionalen Vektorraum über dem Körper mit zwei Elementen $GF(2)$ konstruieren, indem man jede Kante $e = (u, v)$ als $\chi(u) + \chi(v)$, also als den Vektor, der Einsen in Koordinaten u und v und sonst nur Nullen in den übrigen Koordinaten hat. Kreise im Graphen entsprechen dann minimal linear abhängigen Punktmengen in dieser Konfiguration. In dieser Repräsentierung gehen die Knoten anscheinend „verloren“, trotzdem bleibt einiges an Information erhalten. Ein tieferes Verständnis der Problemcharakteristika erhofft man sich von der Untersuchung, welche Aussagen in dem Lineare Algebra Modell noch Gültigkeit haben.

Große Kreise

Das *Rundreiseproblem*, bei dem eine kürzeste Rundreise durch eine Anzahl vorgegebener Städte gesucht ist, ist eine der populärsten Aufgabenstellungen in der kombinatorischen Optimierung. Verwandt damit ist das Problem eines *hamiltonschen Kreises*, das ist ein wiederholungsfreier Kreis, der alle Knoten besucht, in Graphen. Auch dieses Problem ist kombinatorisch schwer und man nimmt an, daß es keine sinnvolle Charakterisierung der Existenz eines solchen Hamiltonkreises gibt. Statt dessen muß man sich mit notwendigen oder hinreichenden Bedingungen begnügen. Die wohl bekannteste solche hinreichende Bedingung ist die folgende von Dirac:

Sei $G = (V, E)$ ein Graph ohne Schleifen und Parallelen und sei an jedem Knoten $v \in V$ die Anzahl der ausgehenden Kanten $d(v) \geq \frac{|V|}{2}$. Dann hat G einen Hamiltonschen Kreis.

Wir untersuchten nun, inwiefern diese Aussage sich allein mit Mitteln der Linearen Algebra formulieren und beweisen läßt. Dazu stellt man zunächst fest, daß die Punkte einen $|V|-1$ dimensionalen Vektorraum aufspannen, man spricht vom Rang $r(M)$ des Matroides M auf der Kantenmenge E. Also können wir $|V|$ durch $r(M)+1$ ersetzen. Um $d(v)$ ersetzen zu können, beobachten wir zunächst, daß jeder Kreis einen Knoten gerade oft trifft. Die Bedingung $d(v) \geq \frac{|V|}{2}$ impliziert darüberhinaus, daß jede nichtleere Kantenmenge, die von jedem Kreis gerade oft getroffen wird, wir sprechen vom orthogonalen Komplement des Kreisraums oder den Kokreisen, mindestens $\frac{|V|}{2}$ Kanten enthält. Damit können wir obigen Satz umformulieren.

Sei M ein einfaches binäres Matroid auf der Menge E und habe jeder Kokreis mindestens $\frac{r(M)+1}{2}$ Elemente. Dann hat M einen Kreis der Größe $r(M)+1$.

Leider ist dieser Satz so falsch. Betrachten wir dazu das Fanodual, das aus folgenden Vektoren besteht.

$$F_7^* = \begin{array}{c} \begin{array}{ccccccc} e_1 & e_2 & e_3 & e_4 & f_1 & f_2 & f_3 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} \end{array}$$

Die Matrix hat vollen Rang 4 und alle Kreise haben 4 Elemente, hingegen habe alle Cokreise 3 oder 4, in jedem Fall jedoch mehr als $\frac{4+1}{2}$ Elemente. Allerdings ist das Fanodual als Übeltäter, wenn es um Verallgemeinerungen von graphentheoretischen Sätzen geht, bekannt. Die Matroide, die diesen Übeltäter nicht enthalten sind eine etwas größere Klasse, als die Matroide, die durch total unimodulare Matrizen definiert sind, welche wiederum alle Graphen umfassen. Es gelang uns nun zu zeigen, daß obiger Satz – und ein paar ähnliche Sätze – in regulären Matroiden gelten. Genauer konnten wir die Gültigkeit in einer etwas größeren Klasse zeigen, wobei der Übeltäter diesmal nicht das Fanodual, sondern die Fanoebene selbst ist.

Sei $M \neq F_7^*$ ein ein einfaches binäres Matroid auf der Menge E ohne F_7-Minor und habe jeder Kokreis mindestens $\frac{r(M)+1}{2}$ Elemente. Dann hat M einen Kreis der Größe $r(M)+1$.

Homomorphieäquivalenz und Max-Flow Min-Cut

Homomorphismen zwischen Matroiden sind Abbildungen mit der Eigenschaft, daß Urbilder abgeschlossener Mengen abgeschlossen sind. Wir studierten die Existenz solcher Homomorphismen, bei denen ein Element festgehalten wir. Zwei Matroide M und M' heißen *homomorphieäquivalent*, wenn es sowohl einen solchen Homomorphismus von M nach M' als auch zurück gibt. Wir stellten fest:

Ein Matroid M ist homomorphiäquivalent zu einem Kreis genau dann, wenn für M die Max-Flow Min-Cut Dualität gilt.

Beschränkt man sich nun auf die Klasse binärer Matroide mit Max-Flow Min-Cut Dualität im Dualen, so gilt für jedes solche Matroid M und jedes $k \in \mathbb{N}$:

Entweder es gibt einen Homomorphismus vom Kreis der Länge k nach M oder es gibt einen Homomorphismus von M in den Kreis der Länge $k+1$.

In allgemeinen Matroiden ist dies nur für $k=1$ richtig. Als Spezialfall enthält diese Aussage den Mengerschen Satz:

Ist $G=(V,E)$ ein Graph und $s,t \in V$, so gilt: Entweder es gibt $k+1$ paarweise kantendisjunkte st-Wege oder es gibt einen st-Schnitt mit höchstens k Kanten.